The GODAE High Resolution Sea Surface Temperature Pilot Project: Data Server Requirements

By

Craig Donlon

Presented at the GODAE Data and Product Server Workshop, Biarritz, France, 12th June, 2002.
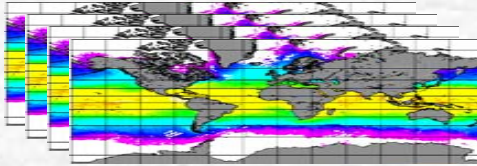
GHRSST-PP Home Page http://ghrsst-pp.jrc.it

The GHRSST-PP Concept

- **In principle, the merging and analysis of complementary satellite and in situ measurements can deliver SST products with enhanced accuracy, spatial and temporal coverage.**
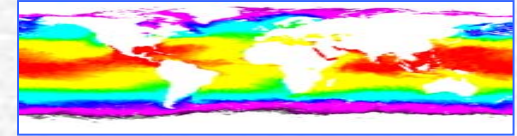

Data Fusion

Quality Control

To facilitate the movement of data

User Applications

GHRSST-PP
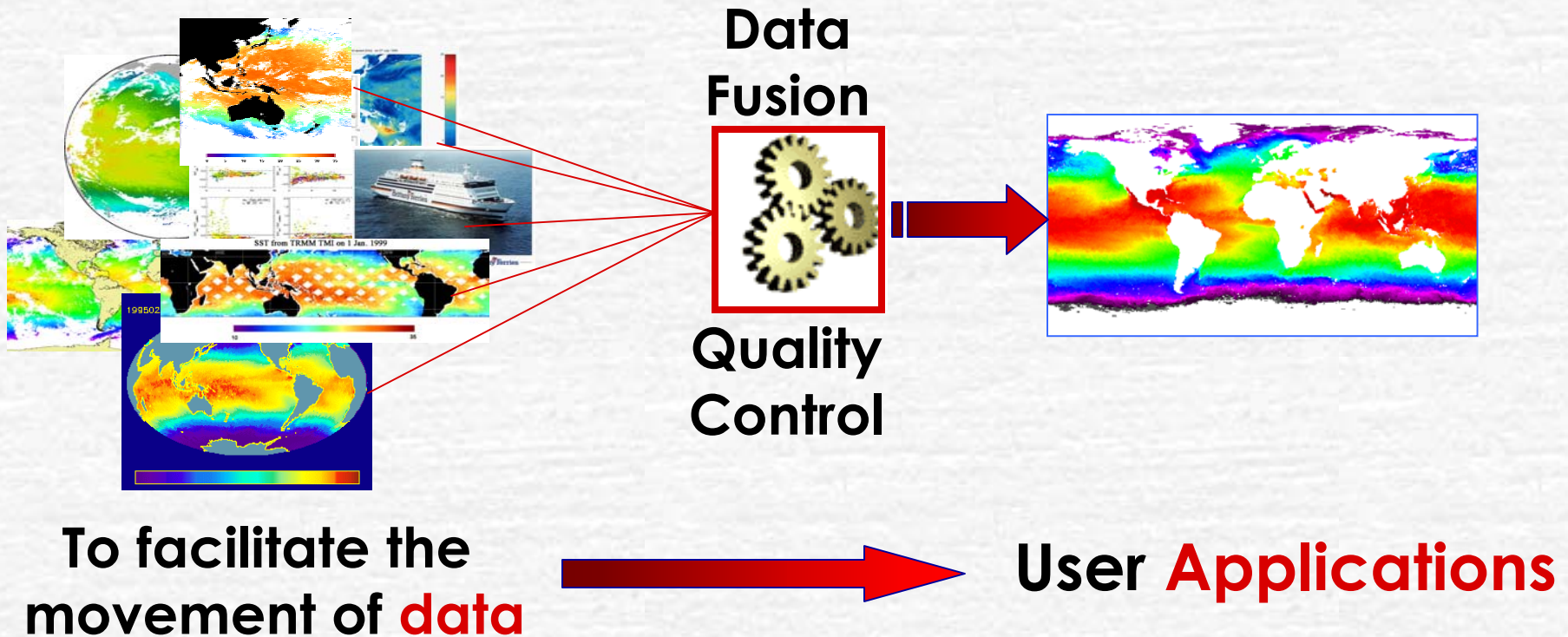
Data Merging and Data Analysis

Merged SST

Analyzed SST

Data Analysis (e.g. OI)

A "stack" or collation of different SST data are calibrated, cleared of cloud and then merged to a common grid. Data are assembled in R/T.	Analyzed data products are derived from the merged "stack".

Merged SST (output every 6 hrs):

- Common grid but variable resolution data
- Individual input data error statistics are preserved
- Products are volatile and change when new data arrives
- Moderate level of quality control
- Only merged 6 hourly data are permanent and archived

Analyzed SST (output every 12 hrs):

- Single output grid together with confidence data (e.g., diurnal signal)
- High level of quality control
- Error statistics include error of source data and of analysis procedure
- Analyzed data are permanent data that are archived.
- Analysed data will be improved via reanalysis

http://www.ghrsst-pp.org

GHRSST-PP

GHRSST-PP Data Products

Characteristic	Merged SST	Analyzed SST
Grid Size	Better than 10 km	Better than 10 km
Temporal resolution	6 hours	12 hours
Accuracy	< 0.5 K absolute 0.1 K relative	< 0.3 K absolute (target) 0.1 K relative
Error statistics	rms. and bias for **each input data stream** at every grid point	rms. and bias for **each output grid point** (no input data statistics are retained)
Coverage	**Regional** **(Best effort Global)**	**Global** **(Regional extracted)**
SST_{skin} product	Yes	Yes
$SST_{sub-skin}$ product	Yes	Yes
SST_{1m} product	Yes	Yes
Cloud mask	**For each input data set**	**Yes**
Confidence data	No	**Yes (sea ice information, diurnal warming mask etc.)**
Nominal data format	Hdf/GRIB/NetCDF	Hdf/GRIB/NetCDF

http://www.ghrsst-pp.org

GoDAE GHRSST-PP

In situ and Satellite Data Integration (ISDI)

GHRSST-PP ST

OPeNDAP,
XML transport

ISDI

(In Situ and Satellite Data integration)

Real time global merging, analysis and validation generating GHRSST-PP products

ISDI-TAG
(ISDI-Technical Advisory Group)

Management of ISDI algorithms, methods and validation

ISDI-DPCF
(Data Product Computational Facility)

Hardware and Software implementation

GHRSST-PP Metadata repository

User Information Services (UIS)
Feedback to ISDI-TAG, RDAC's and GDAC's

Applications and User Services (AUS)
Applications and services monitoring and feedback

*Scheme is identical for both R-ISDI and G-ISDI

GHRSST-PP

Estimated GHRSST-PP Data Volume

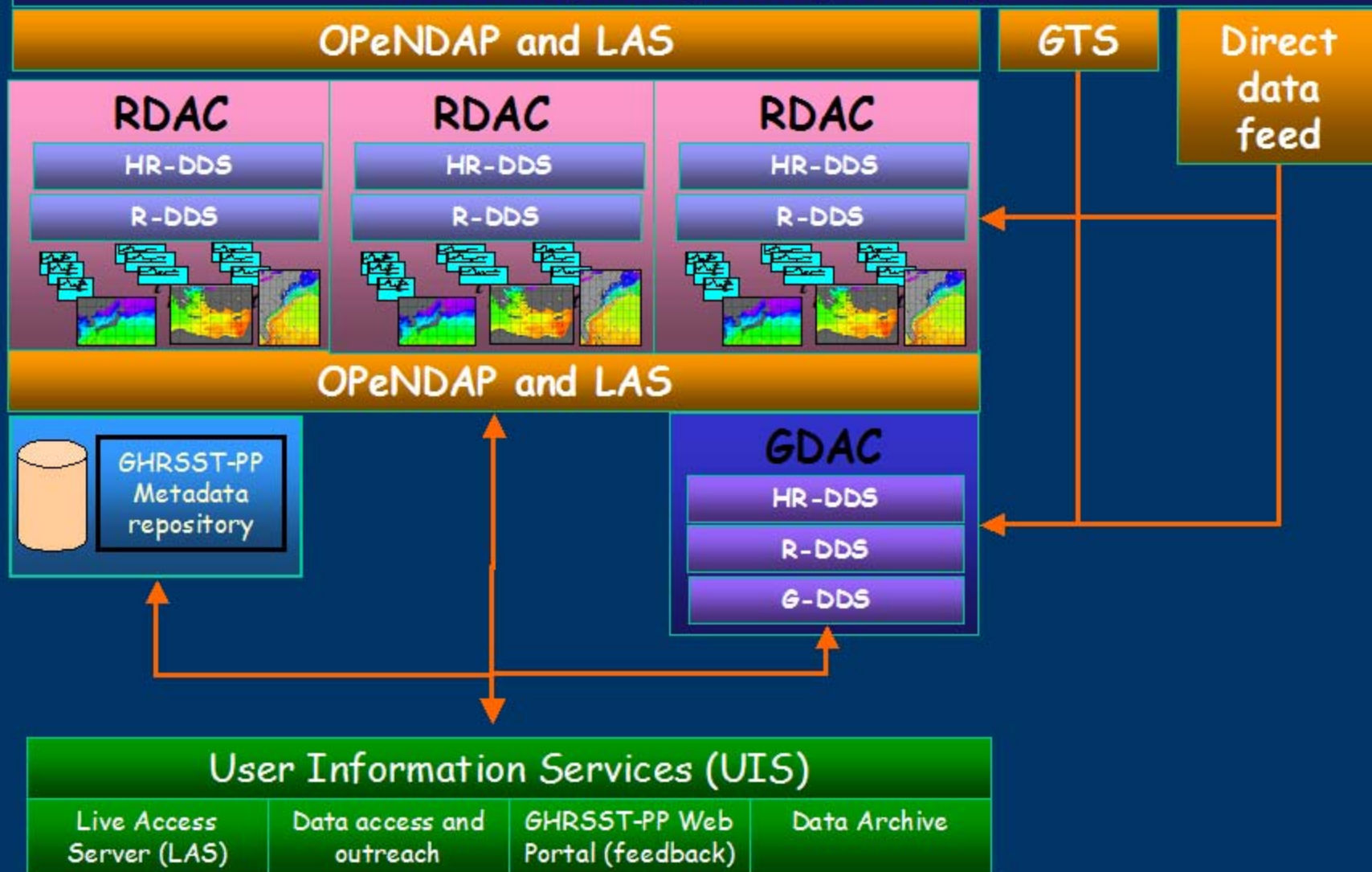| Data set (global)	L1b (multi-ch)	L3 (SST product only)	Frequency	Total (uncompressed)	Retention period
4km (AVHRR, MODIS, GLI)	~2-3Gb	~120Mb	2 x daily	240Mb	24 hrs
18 km (AATSR type)	~2-5Mb	~2Mb	Daily	2Mb	24 hrs
AMSR 25 km	~20Mb	~2Mb	2 x daily	10Mb	24 hrs
TMI (25 km, 8bit)	~3Mb	~1Mb	Daily	1Mb	24 hrs
MSG (~4km)	~0.6Gb	~30Mb	8 x daily	240Mb	6 hrs
GOES (~4km)	~0.6Gb	~30Mb	8 x daily	240Mb	6 hrs
GMS (~4km)	~0.6Gb	~30Mb	8 x daily	240Mb	6 hrs
Buoy & ship	~100Mb	N/A	GTS	100Mb	24 hrs + archive
SSM/I (8bit, wind,sea ice)	~3Mb	~1Mb	Daily	1Mb	24 hrs
Merged GHRSST-pp products	N/A	100-300 Mb	6 hourly	400 – 1.2 Gb(!)	6 hrs + archive
Analyzed GHRSST-pp products	N/A	~250 Mb	12 hourly	500 Mb	12 hrs + archive

GoDAE GHRSST-PP

GHRSST-PP Diagnostic Data Set (DDS)

Distributed In Situ, Satellite and model data sets (SST, wind, solar radiation, fluxes, SOOP, DBP etc)

OPeNDAP and LAS

GTS

Direct data feed

RDAC	RDAC	RDAC
HR-DDS	HR-DDS	HR-DDS
R-DDS	R-DDS	R-DDS

OPeNDAP and LAS

GHRSST-PP Metadata repository

GDAC
- HR-DDS
- R-DDS
- G-DDS

User Information Services (UIS)

Live Access Server (LAS)	Data access and outreach	GHRSST-PP Web Portal (feedback)	Data Archive

The Diagnostic Data Set

- **Ultimately, the GHRSST-PP DDS is the primary GHRSST-PP quality control tool**
  - **Within the project**
  - **Within the AUS**
- **The DDS is a resource for:**
  - **Monitoring of input satellite data streams**
  - **Developing new data merging strategies, tools and methods**
  - **For understanding differences between complementary data**
  - **Validating and monitoring GHRSST-PP data products**
- **The DDS may be implemented as a distributed system linked via the GHRSST-PP metadata repository (e.g., using OPeNDAP and LAS)**

GHRSST-PP

GHRSST-PP Data Movements

- **Types of data movement in GHRSST-PP**
  - **Data provider > RDAC and GDAC** (Dedicated links + GTS)
    - **High volume and frequency (<3 hrs for geostationary feeds)**
  - **RDAC > GDAC** (Dedicated links + GTS)
    - **High volume low frequency (6 hrs for data merging)**
  - **DDS <> RDAC ¦¦ GDAC** (use OPeNDAP and LAS)
    - **Low volume high frequency (whenever available)**
  - **DPCF operations** (external links and visualisation LAS and OPeNDAP)
    - **Movement of data within system but linked to DDS and satellite data holdings**
  - **GHRSST-PP > Users and Applications** (use OPeNDAP and LAS)
    - **Analyzed fields: Low volume, moderate frequency (12 hrs)**
    - **Merged fields: High volume, moderate frequency (6 hrs)**

GHRSST-PP

Main Data Server Issue

- **Merged SST** data place the **highest demands** for the GHRSST-PP
  - Geostationary data sets are large and frequent (can be every hour)
  - Data is required at the GDAC to "patch" the areas not covered by the RDAC
  - Also the data sets most useful to GODAE !!

- **LAS and OPeNDAP will suffice for many operations but dedicated links are required within the GHRSST-PP.**

GHRSST-PP

Data Transport Operations

- **Automatic product subscription push/pull feeds to Applications**
  - Merged and analysed products: high volume, 6hrs
  - What is the general **preferred format** (no single solution)
  - Interoperable tools are require for format conversion
- **Manual and automatic DDS feed/manipulation**
  - OPeNDAP between RDAC & GDAC
  - OPeNDAP access Individual analyses by scientists
  - Direct feed and GTS
- **DPCF operations**
  - Internal to system but linked to data resources at GDAC, RDAC and DDS.
- **Ad hoc data visualisation/ordering/product supply**
  - Low volume, low frequency (6hrs) linked to UIS and AUS
  - LAS (Custom composites and visualisation)

GHRSST-PP

GHRSST-PP Metadata

- **Adhere to metadata standards (..?)**
  - Follow GODAE policy (which is ...?)
  - Should be fully interoperable (e.g., OPeNDAP,GCMD)
  - Should use XML based packaging
- **Metadata will be produced for all data and operations**
  - Ingestion, DDS, products, merging and analysis
  - Metadata will be automatically created and sent to GDAC
  - Source metadata to remain with data
- **Central repository at GDAC and mirrored at RDAC**
  - Master and slave arrangement or simple mirror
  - Timeliness is critical
- **Should be accessible via web and direct access**
  - Web for general user data search at UIS/AUS
  - Direct access for operational access (AUS)
  - Should be a tool and not a burden !

http://www.ghrsst-pp.org

GHRSST-PP

GHRSST-PP data server requirements

- **Need dedicated links for GHRSST-PP GDAC<> data providers and RDAC**
  - High data volume for Merged products
  - Timely data supply
- **DDS as a distributed archive**
  - Use OPeNDAP and LAS
- **UIS at multiple data centers**
  - RDAC, GDAC, PODAAC, JMO etc.
  - Use OPeNDAP and LAS
- **AUS projects linked to UIS interface**
  - Dedicated subscription push feeds as required
  - OPeNDAP and LAS for pull feeds

GoDAE **GHRSST-PP**

GHRSST-PP server status

- **First GHRSST-PP GDAC will be at the US GODAE Monterrey data server**
  - **Linked to GHRSST-PP DPCF**
- **Japanese RDAC**
  - **NASDA AMSR, AMSR/E data server**
  - **NASDA/EORC/JMA/University of Tohoku**
- **European RDAC**
  - **ESA "Medspiration"**

GoDAE **GHRSST-PP**

GHRSST-PP example DDS locations

Proposed GHRSST-PP DDS sites v1

GHRSST-PP ——— SIMBIOS ✳ PIRATA ✳ TAO ✳

GoDAE
GHRSST-PP